

SEARCH ENGINE OPTIMIZATION

Eddie Sparpaglione
Sussex County

Definition

- Process of improving the volume and quality of traffic to a web site from search engines
- Interested in the return and ranking (order) of web site in results

Getting Indexed

- Search engines use **crawlers** to find pages for their results
- Crawlers may look at a number of different factors when crawling a site
 - ▣ Not every page is indexed by the search engines
 - ▣ Distance of pages from the root directory may be a factor

Preventing Indexing

- Avoid undesirable content in search engines
- Webmasters can instruct spiders not to crawl certain files or directories in the **robots.txt** file
 - ▣ When a search engine visits a site, the robots.txt located in the root directory is the first file crawled.
 - ▣ The robots.txt file is then parsed, and will instruct the robot as to which pages are not to be crawled
- Can prevent individual pages from being crawled by using a meta tag for robots in the page header

Robots.txt Explored

- Text file that tells robots what to ignore
- Has to be stored in the root directory as robots.txt
- Examples:

```
User-agent: *  
Disallow: /cgi-bin/  
Disallow: /tmp/  
Disallow: /~joe/
```

*Allows all robots access
to three directories*

```
User-agent: Google  
Disallow:  
  
User-agent: *  
Disallow: /
```

Allow a single robot

```
User-agent: BadBot  
Disallow: /
```

Excludes a single robot

SEO Techniques

White Hat vs. Black Hat

White Hat Techniques

- **Definition:** *Techniques that conform to search engines' guidelines and involve no deception (search engine approved)*
 - Concerned with ensuring that content a search engine indexes and ranks is the same content a user will see
 - Web designers should create content for users and make that content accessible to spiders (do not design for spiders)
-

Black Hat Techniques

- **Definition:** *Techniques to improve rankings that involve deception (not approved by search engines)*
- Examples:
 - Use text that is hidden, either as text colored similar to the background, in an invisible div, or positioned off screen
 - Another method gives a different page depending on whether the page is being requested by a human visitor or a search engine, a technique known as [cloaking](#).
- Search engines may penalize sites they discover using black hat methods, either by reducing their rankings or eliminating their listings from their databases altogether

Listing Order Explained

- A search engine's objective is to present each result in the order of its relevancy to the search query
- Search engine method = *algorithm*
- Most algorithms are kept secret

Search Algorithm Example:

Google's PageRank Algorithm

- Measures the probability that one would end up on a given page when clicking links randomly while surfing.
- If one ends up at a dead end, that is, a page with no links, a new page is selected randomly.
- Pages with higher PageRank are thus considered more important or reliable by Google.
- PageRank can be understood as a democratic process: each page on the web has one vote, and distributes that vote among the other pages it links to. The votes of pages that themselves have higher PageRank are given greater weight in the results.

SEO Tips and Tricks

Practical Methods to Increase Site's Rank

<title> Tag

- The title tag is one of the most important factors in achieving high search engine rankings
- Excellent place for company name
- Title tag should be in the head tag of every page
- Title Tags Should Contain Specific Keyword Phrases
 - ▣ Example: If your company is "Johnson and Smith Inc.," a tax accounting firm in Texas, you shouldn't place only the words "Johnson and Smith Inc." in your title tag, but instead use something like "Johnson and Smith Inc. Tax Accountants in Texas."
 - ▣ Example: Subpages. Also place the titles of subpages after your company title such as "Johnson and Smith Inc. - View Tax Records"

<meta> Tags

- Provide owner information, site description and site keywords
- However due to abuse of these tags, they are mostly ignored by search engines
- Still implement these tags for the rare occasions they are used.

```
<meta name="Author" content="Municipal Web Developers Group" />
```

Author tag tells who owns the site... least important to search engines

```
<meta name="description" content="Municipal Web Developers Group  
Homepage" />
```

Description tag increase page rank for this description and provides a snippet for search engines to display... Most browsers use the first line of text instead

```
<meta name="keywords" content="municipal web developers city town  
delaware" />
```

Give the keywords of the site... Mostly useless due to abuse

Research Keywords

- Your target audience may be searching for something entirely different than the keywords you suggest
- Use a keyword optimization tool to generate relevant keywords
 - <http://www.highrankings.com/keyworddiscovery>
 - <http://www.highrankings.com/wordtracker>
 - <https://adwords.google.com/select/KeywordToolExternal>

Program Your Site to be “Crawler-Friendly”

- Search engines can't fill out forms, can't search your site, can't read JavaScript links and menus, and can't interpret graphics and Flash
 - You Should:
 - ▣ Provide an alternate way to navigate site
 - ▣ Always have some sort of HTML link to other pages (can be graphics)
 - ▣ Label links descriptively, do not use “click here”
-

Increase Links To Your Site

- The number of sites linked to you is critical to search engine optimization
- Make your site full of useful information and the links will come
- Make sure other site's links to you are correct
- Tip:
 - ▣ Type **link:www.yourdomain.com** in a search engine to see how many sites link to you

Sitemaps

- **Definition:** List of the pages on your site
 - ▣ Can provide search engine with URLs it would not be able to get otherwise (not linked)
 - ▣ Created using HTML, text file or XML (preferred)
 - ▣ Should be submitted for each site to improve site ranking

Google Recommendations

What does Google recommend?

Google Tips for Webmasters

Steps to Improve Rating

- Submit site to Google
 - ▣ <http://www.google.com/addurl.html>
- Submit a sitemap
- Make sure all the site that should know about your pages are aware your site is online

Google Tips for Webmasters

- Make a site with a clear hierarchy and text links. Every page should be reachable from at least one static text link.
- Offer a site map to your users with links that point to the important parts of your site. If the site map is larger than 100 or so links, you may want to break the site map into separate pages.
- Create a useful, information-rich site, and write pages that clearly and accurately describe your content.
- Think about the words users would type to find your pages, and make sure that your site actually includes those words within it.

Google Tips for Webmasters

- Try to use text instead of images to display important names, content, or links. The Google crawler doesn't recognize text contained in images.
- Make sure that your <title> elements and alt attributes are descriptive and accurate.
- Check for broken links and correct HTML.
- If you decide to use dynamic pages (i.e., the URL contains a "?" character), be aware that not every search engine spider crawls dynamic pages as well as static pages. It helps to keep the parameters short and the number of them few.
- Keep the links on a given page to a reasonable number (fewer than 100).

Google Resources

Google provides excellent tools for webmasters

Google Resources for Webmasters

- Google's webmaster tools are free
 - Include: Overview, Diagnostics, Statistics, Links, Sitemaps and Tools
 - <http://www.google.com/webmasters/tools/>
 - View crawl statistics, see top search queries, generate robots.txt files and more!
-